It’s hard to believe a year has passed since our last annual report. Time seems to move quickly when you’re engaged in challenging, fulfilling work, as the Research Computing team continues to be. Their contributions over 2016/17 in support of the institution’s research enterprise were many, a number of which we’re pleased to highlight for you in this report.

This year saw the launch of DATASTORE, an institutional service catering to academic data. DataStore removes barriers between researchers and secure storage, providing every faculty member with a reliable home for their digital work. As we rapidly approach a petabyte of allocated storage and make plans to grow beyond, it is comforting to think of the volume of research output now being protected against loss.

Beyond the foundation of secure storage, sharing data enables collaboration, accelerates research, and can unlock entirely new avenues of discovery. In support of this, Research Computing has continued working with our partners on the development of a national-scale service, the Federated Research Data Repository. Set to enter limited production this year, FRDR will aid not only in the preservation of research data, but also in its discoverability, helping to unify the ecosystem of Canadian research data repositories.

Responding to faculty looking to move away from paper-based records, Research Computing conducted an Electronic Lab Notebook pilot over 2016/17. Having demonstrated the usefulness of this software in collecting and accessing data and notes, we are progressing to a campus-wide service for University researchers.

2017/18 will see an increased campus-wide focus on IT security. Research Computing will help our community navigate both technical and policy changes to ensure data is protected from threats, while maintaining research agility and productivity.

I offer my sincerest thanks to this team of talented, committed professionals, and thank you too, for letting us play our part in enabling your research success.

Chad Coller
Director, ICT Academic & Research Technology

The network illustration featured in the background throughout this report is an abstraction of the connectome of the human brain. For Research Computing, it symbolizes the network of collaborators involved in a successful research enterprise.
Who We Are

Enabling research success by providing services and technologies backed by dedicated experts.

Jason Hlady
Manager - Manages Research Computing and is the RDM lead on Compute Canada’s Science Leadership Council.

MSc (Chem)

Ian MacPhedran
Senior Analyst - Provides support and expertise to researchers across the breadth of IT services.

PhD (Eng), PEng

John Costa
Senior Analyst - Provides expertise in health sciences, developing large IT solutions, and IT policy and governance.

MSc (Biomed)

Song Hu
Analyst - Provides Linux support and support in targeted areas, including Geological Sciences and Engineering.

MSc (EE), PEng

Sergiy Stepanenko
ARC Coordinator - coordinates Compute Canada staff and advanced computing services as the WestGrid Site Lead.

MSc (EE)

Juan Zuñiga-Anaya
Advanced Computing Analyst - Supports USask and Compute Canada researchers using advanced research computing.

PhD (CS)

Todd Trann
Lead Developer - Coordinates development effort across Compute Canada and CARL on the FRDR project.

BSc (EE), (CS)

Keith Jeffrey
Senior Analyst - Manages the FRDR project, liaises with researchers, and provides advice on IT policy and governance.

MSc (EE)

Adam McKenzie
RDM Analyst - Supports USask and Compute Canada researchers on research data management, and develops software for FRDR.

MSc (CS)

Mike Winter
Senior Developer - Develops software for FRDR, including analysis and optimization of preservation software.

MSc (CS)

Accomplishments

Enabling research success one milestone at a time.

25 Publications
The research cluster Plato supported at least 20 research publications and 5 conference proceedings in 2016/17.

1 Place for Data
ICT launched DATASSTORE, a new single storage service to ensure that all researchers have access to secure and reliable storage.

520 Bigger and Faster
Advanced Research Computing staff upgraded Plato with 520 additional CPU cores, enabling more research in less time.

74 Trainees
74 Saskatchewan researchers attended Research Computing Summer School in July for training on a variety of advanced computing topics.

23 ELN Pilot
A successful pilot project saw 23 researchers and a Physics senior undergraduate class evaluate an Electronic Notebook system. Read more on page 19.

16 Attendees
Last spring, Research Computing hosted its first workshop on research data management. Future workshops are planned.

700 Terabytes
Over 700 TB of storage space has been allocated on DATASSTORE to over 150 faculty and their collaborators. Read more on page 10-11.

84 Days Saved
Research Computing helped Dr. Lloyd Balbuena (Psychiatry) save 84 days downloading a genomic dataset from the UK Biobank. 11 weeks turned into 10 hours.

FRDR is featured on page 16.

1 FRDR is featured on page 16.
What Can We Do For You?

Enabling research success through services tailored to the research community.

Advanced Research Computing (ARC)
ICT provides infrastructure and support staff to solve problems that are too big to tackle on desktop computers. The WestGrid / Compute Canada staff help users access Canadian ARC resources, design computer systems, and use ARC systems.

Consulting
With a track record of helping researchers design and build solutions, Research Computing experts provide guidance on computational approaches—programming, algorithms, and systems—and directly collaborate with faculty on projects.

Data Collection
REDCap is a web-based data collection application for building and managing online surveys and databases quickly and securely. Tailored to the specific needs of clinical research, REDCap is ready to meet a whole host of data collection requirements.

Visualization & Image Processing
Applying techniques to see data in novel ways using sophisticated visualization software and hardware is critical to the modern research enterprise. From 2D and 3D digital models to 3D printed models, our specialists can help re-imagine your data.

Linux
Our staff have extensive experience with implementing Linux solutions for research. We currently maintain over 35 research workstations across 10 departments, and a significant number of Linux computers supporting the academic mission.

Research Data Management
The exponential growth of research data means that data management has never been more important. Researchers benefit from techniques that keep their data secure, discoverable, and reusable for the future. Research Computing provides expert support and services to help researchers store and manage their data.

Research Applications
ICT provides many specialized research software applications to support research. These include:

- Origin - graphing and data analysis.
- REDCap - Survey and data collection.
- REDCap LTS - For clinical research.
- ArcGIS - mapping and geographic information.
- MATLAB - numerical computing tool for analysis & simulation.
- Mathematica - mathematical symbolic computation.
- Maple - symbolic and numeric computing environment.
- NVivo - qualitative data analysis.

Custom Software
Researchers rely heavily on software to solve difficult problems. While many quality software applications exist, other problems remain unserviced by commercial products. Research Computing provides guidance and analysis for the development of custom software.
Research Computing Training

Summer School

In collaboration with WestGrid, Research Computing hosted a 4-day summer school in July—an intensive bootcamp for researchers interested in expanding their knowledge and skills. Topics included high performance computing, parallel computing, GPU programming, visualization, data management, and practical applications in bioinformatics and material science. The workshop attracted over 70 faculty, staff, and student registrants. Researchers from throughout Saskatchewan were also in attendance, including from the University of Regina, the Saskatoon Cancer Centre, Agriculture and Agri-Food Canada, and the Canadian Light Source. Speakers included Dr. Alex Razoumov from WestGrid, U of S presenters Drs. Matthew Links (Animal Science), Kevin Green, Ray Spiteri (Computer Science), John Tse and Yansun Yao (Physics), and Research Computing experts Mr. Sergiy Stepanenko and Dr. Juan Zuñiga.

Digital Revolution

The digital revolution has touched all aspects of academic life and for researchers, the changes have been dramatic. The tools and techniques for the collection, processing, and analysis of data are more complicated than ever before. At the same time, the ever-accelerating pace of change has pushed researchers to adapt quickly.

"The tools and techniques for the collection, processing, and analysis of data are more complicated than ever before."
Storage for All

If you’ve ever lost or misplaced data before, you understand the value of reliable storage. Until recently, researchers had a few options for storing data—external drives, USB keys, local drives and the cloud. Many storage options lack backups and security, and data can be difficult to share and hard to find over multiple devices. ICT addressed these issues by offering DATASTORE, an institutional storage service dedicated to faculty and their research, teaching and learning, and scholarly work.

Today, DATASTORE is the go-to storage solution for large and small datasets across campus. For example, the cognitive neuroimaging group in Psychology uses DATASTORE to secure their ever-growing output from their fMRI as they explore the brain. Global Institute for Water Security makes extensive use of DATASTORE as they study the world’s water resources. The Molecular Imaging lab in the College of Medicine relies on DATASTORE to host their microPET/SPECT/CT images used in diagnosis and therapy.

IMPACT

Since DATASTORE’s launch March 31st 2017, over 700 TB¹ have been allocated to faculty and more than 250¹ researchers and collaborators now are using DATASTORE, securing over 170TB of research data.

"...it’s fabulous to have a place where I can store data securely and share it with off campus colleagues who are also involved in the project..."

- Faculty survey response

"This is a valuable service to the campus research community and I’m glad to have it..."

- Faculty survey response

¹ As of August 1st, 2017

"DATASTORE... is helping us increase productivity and collaboration between scientists ...

- Branko Zdravkovic, Data Manager, GIWS
Thank you very much for maintaining this cluster, which provides the essential computing power for my research!

- Prof. Yansun Yao, Physics & Engineering Physics

ICT is doing a great job and we appreciate your efforts in keeping the HPC services up.

- Najeeb Khan, PhD Student, Computer Science

A personal thanks to you and the team. I have worked at a few universities and your support is the best I have experienced.

- Dr. Michael Kehoe, Postdoc, GIWS

Research Computing has gone the extra mile in helping us get a new, innovative clinical trial up and running. Without their help we could not have launched this exciting research project.

- Mr. Scott Corley, Director, Clinical Trials Support Unit

Thank you for your support so far in providing computing and data management services to our research groups and GIWS.

— Howard Wheater, Director, Global Institute for Water Security

ICT is doing a great job and we appreciate your efforts in keeping the HPC services up.

- Najeeb Khan, PhD Student, Computer Science

Enabling Research Success
Dr. Yansun Yao

Materials Under Extreme Conditions

Yansun Yao’s research may be described as ‘materials design under extreme conditions’, an emerging area in computational material science. Extreme conditions—for example, high pressure and temperature—offer ways of creating new materials with unprecedented properties, such as superconductivity, super-hardness, and high-energy density. Dr. Yao develops and applies new simulation tools that use supercomputers to predict novel materials and processes under extreme conditions, utilizing quantum mechanical-thermodynamic approaches and multiple heuristic algorithms. These predictions have led to the creation of several new materials in laboratory in the last few years.

Super Results

Yansun Yao uses a supercomputer as his laboratory workbench. After completing his Ph. D. at the University of Saskatchewan, Yansun joined the Physics department as a faculty member in 2015. On his road to receiving the 2016-17 Arts and Science New Scientist Research Award and Physics & Engineering Physics Research Excellence Award, he has done the majority of his computational work using ICT’s research computing cluster, Plato. Since the beginning of 2016, Yansun has run 9,729 jobs on Plato, using 1.67 million hours (191 years) of compute time. His insight, coupled with the computational environment provided by Research Computing, produced 10 peer-reviewed publications in 2016.

“...Yansun has run 9,729 jobs on Plato, using 1.67 million hours (191 years) of compute time.”

“We are sincerely grateful to the Research Computing Team for providing the computational resource and first-class solutions...”

- Professor Yansun Yao, Physics & Engineering Physics
FRDR open beta testing began in June 2017, and limited production with select research groups, including GIWS, is scheduled for Fall 2017. Full production is expected in 2018. The service will be operated by CARL / Portage, making use of existing expertise already in the libraries at institutions across Canada, and will use the processing and storage facilities of Compute Canada.

The beta testing version is available at https://beta.frdr.ca

National Leadership

Researchers from the Global Institute for Water Security (GIWS) will be putting their data into a new national-scale research data repository: the Federated Research Data Repository (FRDR).

For a special issue in the journal Earth System Science Data, the Global Water Futures team needed a place to store published datasets, as well assign unique Digital Object Identifiers (DOIs). GIWS and ICT teamed up to solve this data repository need, while testing the features of the new repository.

FRDR is being developed by a team led out of Research Computing, and includes members throughout Canada. The Canadian Association of Research Libraries and Compute Canada are sponsoring the development with the goal of providing a one-stop-shop that integrates existing and future repositories.

Reproducibility is fundamental to science. FRDR provides a place for Canadian researchers to publish their data, increase its discoverability, and facilitate proper curation and preservation. FRDR will help researchers meet the data publication and reuse expectations of Canadian funding agencies.

"GIWS intends to publish [its] research data in the Federated Research Data Repository (FRDR)...."

— Howard Wheater, Director, Global Institute for Water Security
Secure Computing
A secure computing environment is essential to successful academic and research endeavours. Over the next year, ICT will continue to protect the campus computing environment against cyber threats (e.g. malware, viruses, ransomware) through a number of technical and policy changes. Research Computing will work with the research community to develop solutions that help researchers safeguard their data, maintain productivity, and achieve research success.

Globus Data Sharing
The ability to collaborate and share data between teams and team members is fundamental to research success. This fall, ICT will release Globus—a powerful data management and file sharing service—to help researchers move large datasets effortlessly. The service will be connected to DATASTORE to simplify file sharing, across campus or across the world. Supported by Compute Canada, moving data to the national computing infrastructure will be simple and convenient.

Electronic Notebooks
Electronic Lab Notebooks (ELNs) provide a trusted place to organize and manage research data. They enable collaborative note taking, centralized data collection, searching and discovery, version control, security, and backups. ELNs empower researchers to more easily meet University policy for data retention and research conduct. Building on the success of a pilot involving researchers from several departments earlier this year, Research Computing will launch an ELN service in late 2017.

Equipment Hosting
Researchers encumbered by the operational maintenance of computing infrastructure—servers and storage—in their labs will have the option to have their qualified equipment hosted in one of ICT’s professionally managed data centers.
Professionalism
Excellence
Passion
research_computing@usask.ca